

The local Steiner problem in finite-dimensional normed spaces

Konrad J. Swanepoel

ABSTRACT. We develop a general method for proving that certain star configurations in finite-dimensional normed spaces are Steiner minimal trees. This method generalises the results of Lawlor and Morgan (1994) that could only be applied to differentiable norms. The generalisation uses the subdifferential calculus from convex analysis. We apply this method to two special norms. The first norm, occurring in the work of Cieslik, has unit ball the polar of the difference body of the n -simplex (in dimension 3 this is the rhombic dodecahedron). We determine the maximum degree of a given point in a Steiner minimal tree in this norm. The proof makes essential use of extremal finite set theory. The second norm, occurring in the work of Mark Conger (1989), is the sum of the ℓ_1 -norm and a small multiple of the ℓ_2 norm. For the second norm we determine the maximum degree of a Steiner point.

1. Introduction

There is a vast literature on Steiner minimal trees, mostly in graphs, Hamming space and related word spaces, the Euclidean plane, and the Manhattan plane, with applications in VLSI design [20] and phylogenetics [8]. See e.g. the monographs [7, 18, 19]. Cockayne [9] considered Minkowski planes (two-dimensional normed spaces). Cieslik initiated the study of Steiner minimal trees in general finite-dimensional normed spaces [5, 6]. Recently, other norms besides Euclidean and Manhattan have found applications [4, 22, 23, 24, 25, 26], while the geometry of certain high-dimensional normed spaces is related to the metric spaces found in the mathematical study of phylogenetic trees [8]. These trees are also of interest in differential geometry; see especially the work of Morgan and his students [28, 29, 1, 10, 21].

We denote an n -dimensional normed space or *Minkowski space* $(\mathbb{R}^n, \|\cdot\|)$ by \mathbb{M}^n (see Section 2.3). A *Steiner tree* T of a finite set of points N in \mathbb{M}^n is a tree with vertex set $V = V(T)$ and edge set $E = E(T)$ where $N \subseteq V$. The points in N are called *nodes* or *given points*, and the points in $V \setminus N$ *Steiner points* or *auxiliary*

2000 *Mathematics Subject Classification.* Primary 49Q10; Secondary 05C05, 05D05, 52A21, 52A41, 52B40.

This material is based upon work supported by the South African National Research Foundation under Grant number 2053752.

points. The *length* of a Steiner tree T is

$$\ell(T) := \sum_{\mathbf{x}, \mathbf{y} \in E(T)} \|\mathbf{x} - \mathbf{y}\|.$$

Any finite set of points N in a Minkowski space has at least one tree of minimum length [9]. Such a shortest Steiner tree of N is called a *Steiner minimal tree* (SMT) of N .

A *star* is an SMT such that one of the vertices, called the *centre* of the star, (either a Steiner point or a node) is joined to all the other vertices. The *vertex figure* of a vertex in an SMT is the set of vectors from the centre of the star to the other vertices. Note that only the directions of the vectors matter: If a set of non-zero vectors is a vertex figure in some SMT, then we may replace each vector with another in the same direction, to form a new SMT. In an earlier paper [37] we called the problem of describing the vertex figures of nodes and Steiner points the *local Steiner problem*. Here we give a solution of the local Steiner problem in Minkowski spaces, formulated in terms of norming functionals (Section 4.2).

We denote the maximum degree of a given point in an SMT in \mathbb{M}^n , with the maximum taken over all possible SMTs, by $d(\mathbb{M}^n)$. Let $s(\mathbb{M}^n)$ be the maximum degree of a Steiner point in an SMT in \mathbb{M}^n , where the maximum is taken over all possible SMTs in \mathbb{M}^n . If T is an SMT of N , then T is clearly still an SMT of any N' where $N \subseteq N' \subseteq V(T)$. It follows that $s(\mathbb{M}^n) \leq d(\mathbb{M}^n)$. We use the characterization of vertex figures to obtain information on these values for certain spaces.

Cieslik [5] proved that $d(\mathbb{M}^n)$ is bounded above by the *Hadwiger number* or *translative kissing number* of the unit ball $B(\mathbb{M}^n)$, i.e., the maximum number of mutually non-overlapping translates of $B(\mathbb{M}^n)$ that all touch $B(\mathbb{M}^n)$. This already gives an upper bound depending only on n , since the Hadwiger number is bounded above by $3^n - 1$ [16]. He proved this more generally for all minimal spanning trees, a result rediscovered in [30] for the special case of ℓ_p norms. Cieslik [5], [7, Conjecture 4.3.6] also made the following conjecture:

CONJECTURE 1 (Cieslik [5, 7]). *The maximum degree of a given point in an SMT in any n -dimensional Minkowski space \mathbb{M}^n satisfies*

$$d(\mathbb{M}^n) \leq 2^{n+1} - 2,$$

with equality if and only if \mathbb{M}^n is isometric to the space \mathbb{M}_Z^n which has unit ball $\text{conv}([0, 1]^n \cup [-1, 0]^n)$.

The unit ball of \mathbb{M}_Z^2 is an affine regular hexagon, and of \mathbb{M}_Z^3 an affine rhombic dodecahedron. The norm of \mathbb{M}_Z^n is

$$\|(\mathbf{x}(1), \dots, \mathbf{x}(n))\| = \max_{\mathbf{x}(i) \geq 0} \mathbf{x}(i) - \min_{\mathbf{x}(i) \leq 0} \mathbf{x}(i).$$

See Fig. 1 for two affine representations, one as in the conjecture, and the other with all faces congruent. In Section 2.4 we represent \mathbb{M}_Z^n in a different, more convenient way.

Since any edge joining two points in an SMT can be replaced by a piece-wise linear path consisting of segments parallel to the vectors pointing to the extreme points of the unit ball, we obtain the following well-known lemma.

LEMMA 2. *If the unit ball of \mathbb{M}^n is a polytope with v vertices, then $d(\mathbb{M}^n) \leq v$.*

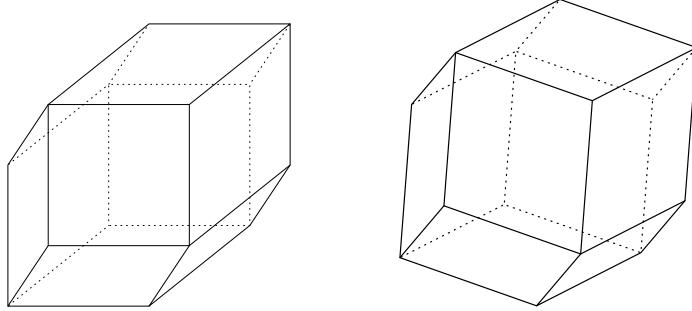


FIGURE 1. Two representations of the rhombic dodecahedron

It follows that $d(\mathbb{M}_Z^n) \leq 2^{n+1} - 2$. Cieslik [6] proved the case $n = 2$ of his conjecture. However, we show that $d(\mathbb{M}_Z^n) < 2^{n+1} - 2$ for all $n \geq 3$, thus partially disproving the conjecture. It is not difficult to see that $d(\mathbb{M}_Z^3) < 14$ by shortening the star connecting the origin to the vertices of the rhombic dodecahedron. If there exists an SMT with a given point of degree 14, then the star joining \mathbf{o} to the vertices of the unit ball will be an SMT of the 14 vertices together with the origin. In Fig. 2 on the left we have three of the fourteen edges from the origin to the vertices. These three vertices, one of degree 4 and two of degree 3, are on the same facet of the unit ball. The length of this subtree is 3. On the right we show that it can be shortened to a length of $5/2$ by introducing two Steiner points. This sub-configuration is therefore forbidden in any star joining the origin to a subset of the vertices, and it easily follows that $d(\mathbb{M}_Z^3) \leq 10$. This turns out to be the correct value (Fig. 3).

THEOREM 3. *For any $n \geq 3$, $d(\mathbb{M}_Z^n) = \binom{n+2}{\lfloor (n+2)/2 \rfloor}$.*

The proof is in Section 5. We also describe all possible vertex figures of a node in an SMT in \mathbb{M}_Z^n . From the proof we obtain the following corollary.

COROLLARY 4. *In the Minkowski space \mathbb{M}_Z^n , the star joining \mathbf{o} to any set A of non-zero points is an SMT of $A \cup \{\mathbf{o}\}$ if and only the star joining \mathbf{o} to any subset B of A of size at most 4 is an SMT of $B \cup \{\mathbf{o}\}$.*

For all $n \geq 2$ there is, up to isometries, a unique configuration attaining the maximum degree (Theorem 23 in Section 5). These spaces give the largest known

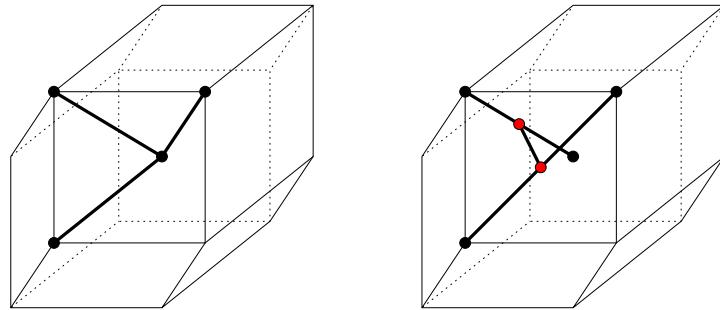


FIGURE 2. Shortening a tree in the rhombic dodecahedral norm

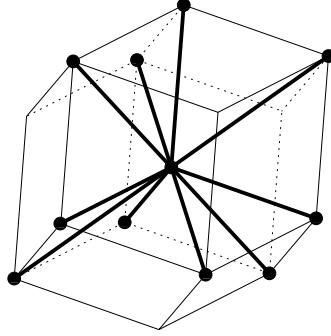


FIGURE 3. A given point of degree 10 in the rhombic dodecahedral norm

degrees of SMTs in Minkowski spaces of dimensions 2 to 6, in fact $d(\mathbb{M}_Z^3) = 10$, $d(\mathbb{M}_Z^4) = 20$, $d(\mathbb{M}_Z^5) = 35$, and $d(\mathbb{M}_Z^6) = 70$. For $n \geq 7$, $d(\ell_\infty^n) = 2^n$ is larger. It is not at all clear whether \mathbb{M}_Z^n maximises $d(\mathbb{M}^n)$.

Morgan [28, p. 42], [29, p. 129] asks the following question:

CONJECTURE 5 (Morgan [28, 29]). *The maximum degree of a Steiner point in an SMT in any n -dimensional Minkowski space \mathbb{M}^n satisfies*

$$s(\mathbb{M}^n) \leq 2^n.$$

The asymptotically best known upper bound for both conjectures is [38]

$$s(\mathbb{M}^n) \leq d(\mathbb{M}^n) \leq c2^n n^2 \log n.$$

It is known that $s(\mathbb{M}^2) \leq 4$ for all planes [37]. There are many two-dimensional spaces attaining $s(\mathbb{M}^2) = 4$. Some piece-wise C^∞ , elliptic examples were discovered in [1]. They are characterised in [37].

We give a lower bound for $s(\mathbb{M}_Z^n)$ that is asymptotically correct up to a factor of 2 by Theorem 3.

THEOREM 6. *For any $n \geq 3$, $s(\mathbb{M}_Z^n) \geq \binom{n+1}{\lfloor (n+1)/2 \rfloor}$. Therefore, $s(\mathbb{M}_Z^n) = \Theta(2^n/\sqrt{n})$.*

The proof is in Section 6.

The sharp upper bound for differentiable norms is $s(\mathbb{M}^n) \leq d(\mathbb{M}^n) \leq n+1$ [21, 36]. It is well-known that $s(\ell_\infty^n) = d(\ell_\infty^n) = 2^n$ [28] and $s(\ell_1^n) = d(\ell_1^n) = 2n$ (see also Section 6). For the ℓ_p norm, $1 < p < \infty$, we have $3 \leq s(\ell_p^n) \leq d(\ell_p^n) \leq 7$ if $p > 2, n \geq 2$, and $\min\{n, \frac{p}{(p-1)\ln 2}\} \leq s(\ell_p^n) \leq d(\ell_p^n) \leq 2^{p/(p-1)}$ if $1 < p < 2, n \geq 3$; see [36], where more detailed estimates are obtained.

Conger [10] showed that $s(\mathbb{R}^3, \|\cdot\|_1 + \lambda \|\cdot\|_2) \geq 6$ for all $0 < \lambda \leq 1$. These norms are piece-wise C^∞ and elliptic. In [1] it is shown that $s(\mathbb{R}^2, \|\cdot\|_1 + \lambda \|\cdot\|_2) = 4$ for all $0 < \lambda \leq 2 + \sqrt{2}$. The value $\lambda = 2 + \sqrt{2}$ is sharp, since it follows from the results in [37] that $s(\mathbb{R}^2, \|\cdot\|_1 + \lambda \|\cdot\|_2) = 3$ for all $\lambda > 2 + \sqrt{2}$. We generalise these results as follows.

THEOREM 7. *Let $\lambda > 0$, and let $\mathbb{M}^n = (\mathbb{R}^n, \|\cdot\|_1 + \lambda \|\cdot\|_2)$. If $\lambda \leq 1$ then $s(\mathbb{M}^n) \leq d(\mathbb{M}^n) \leq 2n$. If $\lambda \leq \sqrt{n}/(\sqrt{n}-1)$ then $d(\mathbb{M}^n) \geq s(\mathbb{M}^n) \geq 2n$.*

The proof is in Section 7. In this regard Conger made the following conjecture [29, p. 128].

CONJECTURE 8 (Conger). *For any piece-wise differentiable, elliptic \mathbb{M}^n we have $s(\mathbb{M}^n) \leq 2n$.*

Our results are based on a characterization of the vertex figures of nodes and of Steiner points in SMTs in arbitrary Minkowski spaces (Section 4). This characterization is found using the subdifferential calculus (Section 3). The characterization of vertex figures of nodes enables us to reduce the determination of $d(\mathbb{M}_Z^n)$ to a purely combinatorial problem in extremal finite set theory (Section 5). The characterization of vertex figures of Steiner points in the case of \mathbb{M}_Z^n is not so easily reducible to combinatorics, hence the partial results of Theorem 6 (Section 6). This situation can be compared to [37], where the characterization of nodes in an arbitrary two-dimensional Minkowski space is much simpler than the characterization of Steiner points. However, the norm $\|\cdot\|_1 + \lambda\|\cdot\|_2$ is sufficiently simple so that the maximum degree of a Steiner point can be determined (Section 7).

In the next section we collect basic definitions used in the paper.

2. Basic definitions

2.1. Signed sets. Let $[n] := \{1, 2, \dots, n\}$, and let $\mathcal{P}[n]$ be the set of all subsets of $[n]$. Let $|A|$ denote the number of elements in the finite set A . We use the following standard notation for signed sets (see [3]). A *signed subset* X of $[n]$ is a vector $X \in \{0, \pm 1\}^n$. We denote the i th component of X by $X(i)$ (reserving subscripts such as X_i for indices). The *positive part* of X is $X^+ := \{i : X(i) = +1\}$, the *negative part* of X is $X^- := \{i : X(i) = -1\}$, the *support* of X is $\underline{X} := X^+ \cup X^-$, and the *zero set* of X is $X^0 := [n] \setminus \underline{X}$. We also write $X = (X^+, X^-)$. Thus any two disjoint subsets A, B of $[n]$ determine a signed set (A, B) . The natural partial ordering on signed sets $X \leq Y$ is defined by $X^+ \subseteq Y^+$ and $X^- \subseteq Y^-$. Two signed sets X, Y are *conformal* if $X^+ \cap Y^- = \emptyset = X^- \cap Y^+$. Denote the *empty signed set* (\emptyset, \emptyset) by \emptyset .

2.2. Vector spaces. We consider \mathbb{R}^n to be the vector space of column vectors $\mathbf{x} = (\mathbf{x}(1), \dots, \mathbf{x}(n))^\top$, with the standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ defined by $\mathbf{e}_i(j) = \delta_{ij}$. The *support* of a vector $\mathbf{x} \in \mathbb{R}^n$ is the signed set $\text{supp}(\mathbf{x}) := (\text{supp}^+(\mathbf{x}), \text{supp}^-(\mathbf{x}))$, where $\text{supp}^+(\mathbf{x}) = \{i : \mathbf{x}(i) > 0\}$ and $\text{supp}^-(\mathbf{x}) = \{i : \mathbf{x}(i) < 0\}$. Let $\mathbf{x}^+, \mathbf{x}^-$ be defined by $\mathbf{x}^+(i) = \max\{\mathbf{x}(i), 0\}$ and $\mathbf{x}^-(i) = \min\{\mathbf{x}(i), 0\}$.

Let $V \subseteq \mathbb{R}^n$ be a subspace, and $\pi : \mathbb{R}^n \rightarrow V$ a projection. Define $\text{supp}(V) := \{\text{supp}(\mathbf{x}) : \mathbf{x} \in V\}$. This is the set of covectors of the oriented matroid of the vector configuration $\{\pi(\mathbf{e}_i) : i = 1, \dots, n\}$ in V ; cf. [3, §2.2].

The *dual* V^* of a vector space V is the set of all linear functionals $\varphi : V \rightarrow \mathbb{R}$. We identify $(\mathbb{R}^n)^*$ with the set of row vectors $\varphi = (\varphi(1), \dots, \varphi(n))$, with standard basis $\epsilon_i = \mathbf{e}_i^\top$. Let V be a subspace of \mathbb{R}^n , and let $\pi : \mathbb{R}^n \rightarrow V$ be an orthogonal projection. The quotient space $\mathbb{R}^n / \ker \pi$ can then be identified with V . The dual of $\mathbb{R}^n / \ker \pi$ is in a canonical way isomorphic to the annihilator of $\ker \pi$, i.e., the subspace $\{\varphi : \ker \pi \subseteq \ker \varphi\}$ of $(\mathbb{R}^n)^*$. We accordingly identify this space with V^* . It is easily seen that with these identifications $\text{supp}(V) = \text{supp}(V^*)$ (as long as π is orthogonal).

For any two vector spaces V and W , $(V \times W)^*$ is canonically isomorphic to $V^* \times W^*$, since a functional $\varphi \in (V \times W)^*$ may be decomposed uniquely as $\varphi(\mathbf{x}, \mathbf{y}) =$

$\psi(\mathbf{x}) + \chi(\mathbf{y})$ for some $\psi \in V^*$ and $\chi \in V^*$; in fact, $\psi(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{o})$ and $\chi(\mathbf{y}) = \varphi(\mathbf{o}, \mathbf{y})$. Conversely, given any $\psi \in V^*, \chi \in W^*$, their sum $\psi(\mathbf{x}) + \chi(\mathbf{y})$ defines a functional on $V \times W$ which we denote (ψ, χ) .

If A and B are subsets of a vector space V , we define their *Minkowski sum* to be $A + B := \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}$.

2.3. Minkowski spaces. A *Minkowski space* is a finite-dimensional vector space V together with a norm $\|\cdot\|$ on V , denoted by $\mathbb{M} = (V, \|\cdot\|)$, or \mathbb{M}^n to indicate that it is n -dimensional. The *unit ball* $B = B(\mathbb{M}) = \{\mathbf{x} \in V : \|\mathbf{x}\| \leq 1\}$ determines the norm uniquely, and we also write $\mathbb{M} = (V, B)$.

The *dual* of $\mathbb{M} = (V, \|\cdot\|)$ is $\mathbb{M}^* = (V^*, \|\cdot\|^*)$, with the dual norm defined by

$$\|\varphi\|^* := \max\{\varphi(\mathbf{x}) : \|\mathbf{x}\| = 1\}.$$

Denote the dual unit ball by B^* . It is well-known that the dual of the dual is again the original space.

The ℓ_p -norm on \mathbb{R}^n is defined by

$$\|(\mathbf{x}(1), \dots, \mathbf{x}(n))\|_p := \left(\sum_{i=1}^n |\mathbf{x}(i)|^p \right)^{1/p}$$

if $p \geq 1$, and

$$\|(\mathbf{x}(1), \dots, \mathbf{x}(n))\|_\infty := \max\{\mathbf{x}(i) : i \in [n]\}.$$

The space ℓ_p^n is $(\mathbb{R}^n, \|\cdot\|_p) = (\mathbb{R}^n, B_p^n)$. The n -cube is B_∞^n , and the n -dimensional cross polytope is B_1^n . The spaces ℓ_∞^n and ℓ_1^n are dual.

By the (Hahn-Banach) separation theorem, each non-zero $\mathbf{x} \in \mathbb{M}^n$ has a *norming functional*, i.e. a functional φ such that $\|\varphi\|^* = 1$ and $\varphi(\mathbf{x}) = \|\mathbf{x}\|$. We denote the set of norming functionals of a non-zero \mathbf{x} by $\partial\mathbf{x}$. Each $\partial\mathbf{x}$ is an *exposed face* of B^* , i.e., the intersection of B^* by some supporting hyperplane, in fact $\partial\mathbf{x} = B^* \cap \{\varphi : \varphi(\mathbf{x}) = \|\mathbf{x}\|\}$. A norm is differentiable if and only if $\partial\mathbf{x}$ is a singleton for each $\mathbf{x} \neq \mathbf{o}$ (and then the norm is C^1).

A norm is *strictly convex* if for all linearly independent $\mathbf{x}, \mathbf{y} \in \mathbb{M}^n$ we have $\|\mathbf{x} + \mathbf{y}\| < \|\mathbf{x}\| + \|\mathbf{y}\|$. This is equivalent to there being no straight segment in the boundary of the unit ball (i.e., *strict convexity* of the unit ball). A norm is strictly convex if and only if the dual norm is differentiable.

A norm $\|\cdot\|$ on V is *elliptic* if $\|\cdot\| - \|\cdot\|_2$ is still a norm for some Euclidean norm $\|\cdot\|_2$ (i.e. for any identification of V with \mathbb{R}^n). We also say that the unit ball of an elliptic norm is *elliptic*. If the norm is C^2 , this is equivalent to requiring that the boundary of the unit ball has positive inward curvature bounded away from 0, where the curvature is defined by any Euclidean structure on the space. A norm is elliptic if and only if its dual is $C^{1,1}$. See [27, §3] for more on elliptic norms.

Note that in functional analysis a differentiable norm is also called a smooth norm, while in differential geometry the word smooth is usually reserved for C^∞ . In differential geometry the unit ball of an elliptic norm is also called uniformly convex, which again has another meaning in Banach space geometry. We therefore avoid the terms smooth and uniformly convex in this paper, and stick to the terminology of the previous paragraph.

If V is a subspace of $\mathbb{M} = (W, \|\cdot\|)$, then the quotient space \mathbb{M}/V is the space W/V with norm $\|\mathbf{x} + V\| := \inf\{\|\mathbf{x} + \mathbf{v}\| : \mathbf{v} \in V\}$. If V is a subspace of $\mathbb{M} = (\mathbb{R}^n, \|\cdot\|)$, a concrete presentation of \mathbb{M}/V can be obtained by using the orthogonal projection $\pi : \mathbb{R}^n \rightarrow V^\perp$. Then \mathbb{M}/V is isometrically isomorphic to $(V^\perp, \pi(B))$,

i.e., we project the unit ball B orthogonally onto V^\perp to obtain the norm. The dual $(\mathbb{M}/V)^*$ is as before the annihilator of $\ker \pi = V$, and the dual norm is a restriction of the dual norm $\|\cdot\|^*$ of \mathbb{M}^* , i.e., the dual is a subspace of \mathbb{M}^* , with the embedding given by the adjoint π^* of the projection π .

2.4. Zonotopes as unit balls. Denote the closed segment from \mathbf{a} to \mathbf{b} by $[\mathbf{a}, \mathbf{b}]$. A *zonotope* is a Minkowski sum of closed segments, equivalently, a zonotope is the projection of an n -cube onto a subspace. The zonotope *determined by* a subspace V of \mathbb{R}^n is the orthogonal projection $Z = \pi(B_\infty^n)$ of the n -cube onto V . As explained in Section 2.3, (V, Z) is isometric to the quotient $\ell_\infty^n / \ker \pi$, and its dual (V^*, Z^*) is isometric to the subspace $\text{im } \pi^*$ of ℓ_1^n . As before, we identify V^* with this subspace. Then $Z^* = V^* \cap B_1^n$. We need a description of the faces of Z^* . Each proper face of Z^* is the intersection of V^* with a proper face of B_1^n , i.e., there is one for each signed set $X \in \text{supp}(V^*)$, namely

$$F'(X) := \{\varphi \in V^* : \|\varphi\|_1 = 1, \text{supp}(\varphi) \leq X\},$$

of dimension $|\underline{X}| - 1$. Since $X \mapsto F'(X)$ is order-preserving, we have that $F'(X)$ is a facet if and only if X is a maximal covector in $\text{supp}(V^*)$. Two faces $F'(X)$ and $F'(Y)$ belong to a common facet if and only if X and Y are conformal.

We mention that the corresponding non-empty faces of Z are indexed by the same set $\text{supp}(V) = \text{supp}(V^*)$. Each such face equals some

$$F(X) := \sum_{i \in X^+} \pi(\mathbf{e}_i) - \sum_{j \in X^-} \pi(\mathbf{e}_j) + \sum_{k \in X^0} [-\pi(\mathbf{e}_k), \pi(\mathbf{e}_k)],$$

of dimension $|X^0|$ (see [3, §2.2]). The mapping $X \mapsto F(X)$ is order-reversing, and so $F(X)$ is a vertex if and only if X is a maximal covector in $\text{supp}(V^*)$, while two faces $F(X)$ and $F(Y)$ intersect if and only if X and Y are conformal.

We now consider the Minkowski space \mathbb{M}_Z^n defined in the introduction. Let π be the orthogonal projection of \mathbb{R}^{n+1} onto the hyperplane

$$H := \{\mathbf{x} \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} \mathbf{x}(i) = 0\}.$$

This projection is along the vector $\mathbf{j} := (1, 1, \dots, 1)^\top$. Thus \mathbb{M}_Z^n is the quotient $\ell_\infty^{n+1} / \ker \pi$, which we identify with the n -dimensional space H with unit ball $Z_n := \pi(B_\infty^{n+1})$. It is easily seen that the norm is given by

$$\|\mathbf{x}\|_Z := \frac{1}{2} (\max_{i \in [n+1]} \mathbf{x}(i) - \min_{i \in [n+1]} \mathbf{x}(i)).$$

The dual space $\mathbb{M}_Z^n := (\mathbb{M}_Z^n)^*$ is the subspace

$$H^* := \{\varphi \in (\mathbb{R}^{n+1})^* : \sum_{i=1}^{n+1} \varphi(\mathbf{e}_i) = 0\}$$

of ℓ_1^{n+1} , with support

$$\mathcal{Z}_n := \text{supp}(H^*) = \{X \in \{0, \pm 1\}^{n+1} : X^+ \neq \emptyset, X^- \neq \emptyset\} \cup \{\emptyset\}.$$

Note that $F'(\emptyset) = \emptyset$, and all other $F'(X)$ are non-empty proper faces of Z_n^* . The facets of Z_n^* are all the $F'(X)$ with $X \in \{\pm 1\}^{n+1}$, $X^\pm \neq \emptyset$. In general,

$$F'(X) = \{\varphi \in (\mathbb{R}^{n+1})^* : \sum_{i \in X^+} \varphi(\mathbf{e}_i) = \frac{1}{2}, \sum_{i \in X^-} \varphi(\mathbf{e}_i) = -\frac{1}{2}, \text{supp}(\varphi) \leq X\}.$$

Also define $\overline{\mathcal{Z}_n} := \mathcal{Z}_n \cup \{\mathbf{1}\}$, with $\mathbf{1}$ a symbol corresponding to the improper face Z_n^* . We consider $\mathbf{1} \geq X$ for all $X \in \mathcal{Z}_n$. Thus $(\overline{\mathcal{Z}_n}, \leq)$ is the face lattice of Z_n^* .

It is easy to see that Z_n^* is the difference body of the regular n -simplex

$$\Delta_n := \text{conv}\{\epsilon_i : i \in [n+1]\}.$$

Thus Z_2^* is a regular hexagon, and Z_3^* a cuboctahedron.

To show that Z_n is affinely equivalent to $\text{conv}([0, 1]^n \cup [-1, 0]^n)$, let π' be the (non-orthogonal) projection of \mathbb{R}^{n+1} onto the hyperplane

$$H' = \{\mathbf{x} \in \mathbb{R}^{n+1} : \mathbf{x}(n+1) = 0\},$$

again along the vector \mathbf{j} . Then $\pi'|_{H'}$ and $\pi|_H$ are inverses of each other taking Z_n to $\pi'(C^{n+1})$ and vice versa. Identifying H' with \mathbb{R}^n in the obvious way, it is easily seen that $\pi'(C^{n+1}) = \text{conv}([0, 2]^n \cup [-2, 0]^n)$.

3. Subdifferential calculus

In this section we summarise the basic results of the subdifferential calculus needed in this paper. For more on subdifferential calculus, see [31] and [40]. We only consider convex functions defined everywhere in finite-dimensional spaces, so many of the arguments simplify. Although these results are all well-known in convex analysis, they are not easy to locate in the same reference, and the proofs given are usually based on a variety of other more general theorems. For the convenience of the reader we give direct proofs of these lemmas. We use the following version of the (weak) separation theorem [31, Theorem 11.3].

SEPARATION THEOREM. *Let C_1 and C_2 be non-empty convex sets with disjoint relative interiors in a finite-dimensional real vector space. Then there exists a hyperplane H such that C_1 and C_2 are in opposite closed half spaces bounded by H .*

Let V be any Minkowski space with dual V^* . As usual, a function $f : V \rightarrow \mathbb{R}$ is *convex* if $f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in V$ and $0 \leq \lambda \leq 1$. Convex functions on V are continuous on V . A functional $\varphi \in V^*$ is a *subgradient* of a convex function f at the point \mathbf{a} if for all $\mathbf{x} \in V$,

$$f(\mathbf{x}) - f(\mathbf{a}) \geq \varphi(\mathbf{x} - \mathbf{a}).$$

In particular, $\varphi \in V^*$ is a subgradient of f at \mathbf{a} if and only if f attains its minimum value at \mathbf{a} . The *subdifferential* of f at \mathbf{a} , denoted by $\partial f(\mathbf{a})$, is the set of all subgradients of f at \mathbf{a} . Note that for any $\varphi \in V^*$ and $c \in \mathbb{R}$, $\partial(\varphi + c)(\mathbf{a}) = \{\varphi\}$ for all $\mathbf{a} \in V$.

LEMMA 9. *The subdifferential of a convex function $f : V \rightarrow \mathbb{R}$ at any $\mathbf{a} \in V$ is non-empty, compact, and convex.*

See [31, Theorem 24.7].

PROOF. Since $\partial f(\mathbf{a}) = \bigcap_{\mathbf{x} \in V} \{\varphi \in V^* : \varphi(\mathbf{x} - \mathbf{a}) \leq f(\mathbf{x}) - f(\mathbf{a})\}$, which is an intersection of closed half spaces in V^* , we have that $\partial f(\mathbf{a})$ is closed and convex. If $\partial f(\mathbf{a})$ is unbounded, a compactness argument shows that it contains a ray $\{\varphi_0 + \lambda\varphi_1 : \lambda \geq 0\}$. Therefore, $f(\mathbf{x}) - f(\mathbf{a}) \geq \varphi_0(\mathbf{x} - \mathbf{a}) + \lambda\varphi_1(\mathbf{x} - \mathbf{a})$ for all $\mathbf{x} \in V$ and $\lambda \geq 0$. It follows that $\varphi_1(\mathbf{x} - \mathbf{a}) \leq 0$ for all $\mathbf{x} \in V$, a contradiction.

To show that $\partial f(\mathbf{a}) \neq \emptyset$ we use the separation theorem to obtain a hyperplane that separates the point $(\mathbf{a}, f(\mathbf{a}))$ and the *epigraph* of f

$$\{(\mathbf{x}, y) : \mathbf{x} \in V, y \geq f(\mathbf{x})\} \subseteq V \times \mathbb{R}.$$

Thus there exists a non-zero functional $(\varphi, r) \in (V \times \mathbb{R})^* = V^* \times \mathbb{R}$ such that $(\varphi, r)(\mathbf{x}, y) \geq (\varphi, r)(\mathbf{a}, f(\mathbf{a}))$ for all $\mathbf{x} \in V$ and $y \geq f(\mathbf{x})$. Then $\varphi(\mathbf{x} - \mathbf{a}) \geq r(f(\mathbf{a}) - y)$, which gives that $r \geq 0$, since y can be arbitrarily large. If $r = 0$ this also gives $\varphi = o$, contradicting $(\varphi, r) \neq (o, 0)$. Therefore, $r > 0$, and $-\frac{1}{r}\varphi(\mathbf{x} - \mathbf{a}) \leq f(\mathbf{x}) - f(\mathbf{a})$ for all $\mathbf{x} \in V$, giving $-\frac{1}{r}\varphi \in \partial f(\mathbf{a})$. \square

LEMMA 10 (Moreau-Rockafellar Theorem). *Let $f_i : V \rightarrow \mathbb{R}$, $i = 1, \dots, k$ be convex functions. Then for all $\mathbf{a} \in V$,*

$$\partial\left(\sum_{i=1}^k f_i\right)(\mathbf{a}) = \sum_{i=1}^k \partial f_i(\mathbf{a}),$$

where the sum on the right is Minkowski addition.

See [31, Theorem 23.8] and [40, Theorem 2.8.7].

PROOF. The “ \supseteq ”-inclusion is straightforward, but the “ \subseteq ”-inclusion needs the separation theorem (see the second proof on p. 224 of [31]). We prove the case $k = 2$, with the general case following by induction.

Let f and g be convex functions on V and let $\varphi \in \partial(f + g)(\mathbf{a})$. Then

$$f(\mathbf{x}) - f(\mathbf{a}) + g(\mathbf{x}) - g(\mathbf{a}) \geq \varphi(\mathbf{x} - \mathbf{a}) \text{ for all } \mathbf{x} \in V. \quad (1)$$

The sets

$$C_1 = \{(\mathbf{x}, r) \in V \times \mathbb{R} : f(\mathbf{x}) - f(\mathbf{a}) - \varphi(\mathbf{x} - \mathbf{a}) \leq r\}$$

and

$$C_2 = \{(\mathbf{x}, r) \in V \times \mathbb{R} : g(\mathbf{x}) - g(\mathbf{a}) \leq -r\}$$

are both convex and closed with non-empty interior, and by (1) their interiors are disjoint. They have a common point $(\mathbf{a}, 0)$. By the separation theorem there exists a non-zero $(\psi, s) \in (V \times \mathbb{R})^* = V^* \times \mathbb{R}$ such that $(\psi, s)(\mathbf{x}, r) \geq (\psi, s)(\mathbf{a}, 0)$ for all $(\mathbf{x}, r) \in C_1$ and $(\psi, s)(\mathbf{x}, r) \leq (\psi, s)(\mathbf{a}, 0)$ for all $(\mathbf{x}, r) \in C_2$. This gives that for all $\mathbf{x} \in V$,

$$\psi(\mathbf{x} - \mathbf{a}) \geq -sr \text{ for all } r \geq f(\mathbf{x}) - f(\mathbf{a}) - \varphi(\mathbf{x} - \mathbf{a}), \quad (2)$$

and

$$\psi(\mathbf{x} - \mathbf{a}) \leq -sr \text{ for all } r \leq g(\mathbf{a}) - g(\mathbf{x}).$$

Since r can be arbitrarily large in (2), $s \geq 0$. If $s = 0$, then $\psi(\mathbf{x} - \mathbf{a}) = 0$ for all $\mathbf{x} \in V$, contradicting $(\psi, s) \neq (o, 0)$. Therefore, $s > 0$, and we obtain

$$-\frac{1}{s}\psi(\mathbf{x} - \mathbf{a}) \leq f(\mathbf{x}) - f(\mathbf{a}) - \varphi(\mathbf{x} - \mathbf{a})$$

and

$$\frac{1}{s}\psi(\mathbf{x} - \mathbf{a}) \leq g(\mathbf{x}) - g(\mathbf{a}).$$

This gives $\varphi - \frac{1}{s}\psi \in \partial f(\mathbf{a})$ and $\frac{1}{s}\psi \in \partial g(\mathbf{a})$. Adding, we obtain $\varphi \in \partial f(\mathbf{a}) + \partial g(\mathbf{a})$, as required. \square

LEMMA 11. *Let $f_i : V_i \rightarrow \mathbb{R}$, $i = 1, \dots, k$ be convex, and define $f : V_1 \times \dots \times V_k \rightarrow \mathbb{R}$ by $f(\mathbf{x}_1, \dots, \mathbf{x}_k) = \sum_{i=1}^k f_i(\mathbf{x}_i)$. Then the subdifferential of the convex function f is the Cartesian product*

$$\partial f(\mathbf{a}_1, \dots, \mathbf{a}_k) = \prod_{i=1}^k \partial f_i(\mathbf{a}_i) \subseteq V_1^* \times \dots \times V_k^*.$$

See [40, Corollary 2.4.5].

PROOF. By induction it is sufficient to consider the case $k = 2$. Let $f : V \rightarrow \mathbb{R}$ and $g : W \rightarrow \mathbb{R}$ be convex, define $F : V \times W \rightarrow \mathbb{R}$ by $F(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + g(\mathbf{y})$, and choose $\mathbf{a} \in V$, $\mathbf{b} \in W$.

First choose $\varphi \in \partial f(\mathbf{a})$ and $\psi \in \partial g(\mathbf{b})$. Then $f(\mathbf{x}) - f(\mathbf{a}) \geq \varphi(\mathbf{x} - \mathbf{a})$ for all $\mathbf{a} \in V$, and $g(\mathbf{y}) - g(\mathbf{b}) \geq \psi(\mathbf{y} - \mathbf{b})$ for all $\mathbf{b} \in W$. Adding, we obtain $F(\mathbf{x}, \mathbf{y}) - F(\mathbf{a}, \mathbf{b}) \geq (\varphi, \psi)(\mathbf{x}, \mathbf{y}) - (\varphi, \psi)(\mathbf{a}, \mathbf{b})$, which gives $(\varphi, \psi) \in \partial F(\mathbf{a}, \mathbf{b})$.

Now choose $(\varphi, \psi) \in \partial F(\mathbf{a}, \mathbf{b})$. This means that for all $\mathbf{x} \in V$ and $\mathbf{y} \in W$,

$$\varphi(\mathbf{x} - \mathbf{a}) + \psi(\mathbf{y} - \mathbf{b}) \leq f(\mathbf{x}) - f(\mathbf{a}) + g(\mathbf{y}) - g(\mathbf{b}).$$

Setting $\mathbf{y} = \mathbf{b}$ we obtain $\varphi \in \partial f(\mathbf{a})$, and setting $\mathbf{x} = \mathbf{a}$ gives $\psi \in \partial g(\mathbf{b})$. \square

We have already defined the set of norming functionals $\partial \mathbf{x}$ of a non-zero $\mathbf{x} \in \mathbb{M}$. We now extend this notation by defining $\partial \mathbf{o} = B(\mathbb{M}^*)$. The reason is the following:

LEMMA 12. *The subdifferential of the norm of \mathbb{M} at \mathbf{a} is given by $\partial \|\mathbf{a}\| = \partial \mathbf{a}$.*

See [40, Corollary 2.4.16].

PROOF. By definition, $\varphi \in B(\mathbb{M}^*)$ if and only if $\varphi(\mathbf{x}) \leq \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{M}$, which is equivalent to $\varphi \in \partial \|\mathbf{a}\|$. This takes care of the case $\mathbf{a} = \mathbf{o}$.

If $\mathbf{a} \neq \mathbf{o}$, then for any $\varphi \in \partial \mathbf{a}$ and any $\mathbf{x} \in \mathbb{M}$,

$$\varphi(\mathbf{x} - \mathbf{a}) = \varphi(\mathbf{x}) - \|\mathbf{a}\| \leq \|\mathbf{x}\| - \|\mathbf{a}\|,$$

hence $\varphi \in \partial \|\mathbf{a}\|$. Conversely, suppose

$$\|\mathbf{x}\| - \|\mathbf{a}\| \geq \varphi(\mathbf{x} - \mathbf{a}) \text{ for all } \mathbf{x} \in \mathbb{M}. \quad (3)$$

Setting $\mathbf{x} = \mathbf{o}$ in (3) we obtain $\varphi(\mathbf{a}) \geq \|\mathbf{a}\|$, giving $\|\varphi\|^* \geq 1$. Setting $\mathbf{x} = 2\mathbf{a}$ in (3) we obtain $\|\mathbf{a}\| \geq \varphi(\mathbf{a})$, giving $\varphi(\mathbf{a}) = \|\mathbf{a}\|$. Now (3) gives $\|\mathbf{x}\| \geq \varphi(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{M}$, implying $\|\varphi\|^* = 1$. It follows that φ is a norming functional of \mathbf{a} . \square

The distance function $\rho(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$ is easily seen to be convex on $\mathbb{M} \times \mathbb{M}$ by the triangle inequality.

LEMMA 13. *For any $\mathbf{a}, \mathbf{b} \in \mathbb{M}$,*

$$\partial \rho(\mathbf{a}, \mathbf{b}) = \{(\varphi, -\varphi) : \varphi \in \partial(\mathbf{a} - \mathbf{b})\} \subseteq \mathbb{M}^* \times \mathbb{M}^*.$$

PROOF. Let $\varphi \in \partial(\mathbf{a} - \mathbf{b})$. By Lemma 12 we have for any $\mathbf{x}, \mathbf{y} \in \mathbb{M}$,

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\| - \|\mathbf{a} - \mathbf{b}\| &\geq \varphi(\mathbf{x} - \mathbf{y} - \mathbf{a} + \mathbf{b}) \\ &= \varphi(\mathbf{x} - \mathbf{a}) - \varphi(\mathbf{y} - \mathbf{b}) = (\varphi, -\varphi)(\mathbf{x} - \mathbf{a}, \mathbf{y} - \mathbf{b}), \end{aligned}$$

which gives $(\varphi, -\varphi) \in \partial \rho(\mathbf{a}, \mathbf{b})$.

Conversely, let $(\varphi, \psi) \in \partial \rho(\mathbf{a}, \mathbf{b})$. Then, in particular, for any $\mathbf{x} \in \mathbb{M}$,

$$\rho(\mathbf{x}, \mathbf{x}) - \rho(\mathbf{a}, \mathbf{b}) \geq \varphi(\mathbf{x} - \mathbf{a}) + \psi(\mathbf{x} - \mathbf{b}),$$

i.e., $(\varphi + \psi)(\mathbf{x}) \leq \varphi(\mathbf{a}) + \psi(\mathbf{b}) - \|\mathbf{a} - \mathbf{b}\|$. Since the right-hand side is independent of \mathbf{x} , we obtain $\varphi + \psi = \mathbf{o}$. \square

4. The local Steiner problem in Minkowski spaces

Our characterization of vertex figures is in terms of so-called reduced Minkowski addition defined on subsets of the dual of a Minkowski space.

4.1. Reduced Minkowski addition. We define the *reduced Minkowski sum* of two closed, convex subsets C and D of the unit ball of $(\mathbb{M}^n)^*$ to be

$$C \boxplus D = \{\varphi + \psi : \varphi \in C, \psi \in D, \|\varphi + \psi\|^* \leq 1\},$$

i.e. $C \boxplus D$ is the usual Minkowski sum $C + D$ intersected by the unit ball B^* of $(\mathbb{M}^n)^*$. Reduced Minkowski addition is clearly commutative but not associative, although it satisfies a weak associative law (see Proposition 14 below). We now consider some elementary properties of this binary operation. Consider a finite non-empty family $\Sigma = \{A_i : i \in I\}$ of operands, where $I \subseteq [k]$ and each A_i is a closed convex subset of B^* . We call I the *support* of Σ . A *parenthesization* of Σ is defined to be a parenthesization, in the usual sense, of some ordering

$$A_{j(1)} \boxplus \dots \boxplus A_{j(i)} \tag{4}$$

of Σ , where $j : [|I|] \rightarrow I$ is some bijection. Stanley calls this a *binary set bracketing* of Σ [35, p. 178], and its enumeration is called Schröder's third problem [33]. We denote a parenthesization of Σ by $\langle \Sigma \rangle$, and we also call the family I the *support* of $\langle \Sigma \rangle$. We may define a parenthesization of Σ recursively as follows.

- If Σ is the singleton $\{A_i\}$, then $\langle \Sigma \rangle = A_i$ is the only parenthesization of Σ .
- If $\Sigma = \Sigma_1 \cup \Sigma_2$, with Σ_1 and Σ_2 disjoint, and $\langle \Sigma_i \rangle$ is a parenthesization of Σ_i for each $i = 1, 2$, then both $\langle \Sigma \rangle = (\langle \Sigma_1 \rangle \boxplus \langle \Sigma_2 \rangle)$ and $\langle \Sigma \rangle = (\langle \Sigma_2 \rangle \boxplus \langle \Sigma_1 \rangle)$ are *equivalent* parenthesizations of Σ .

More generally, two parenthesizations of Σ are *equivalent* if they can be transformed into each other using the commutative law on any subexpression. This is an equivalence relation, and two equivalent parenthesizations clearly evaluate to the same set. The collection of equivalence classes of parenthesizations of a set Σ with support I corresponds bijectively with the collection of abstract trees with $|I| + 1$ leaves labelled by the elements of $\{0\} \cup I$, and with $|I| - 1$ internal vertices of degree 3. We call such a tree a *rooted abstract Steiner tree* on I . The node 0 is the root of the tree, node i corresponds to A_i for each $i \in I$, and each internal vertex corresponds to an instance of \boxplus in the parenthesization corresponding to the tree. Note that we do not distinguish between left and right branches (as is done in the well-known bijection between parenthesizations of ordered expressions and planted, trivalent plane trees, both counted by the Catalan numbers; see [35, Exercise 6.19(b),(f)]). Stanley calls this a *binary total partition* [35, Example 5.2.6] of the set I . The explicit construction of the bijection for our variant is the following. We denote the rooted abstract Steiner tree associated to $\langle \Sigma \rangle$ by $T_0 \langle \Sigma \rangle$.

- If $\langle \Sigma \rangle = A_i$ then $T_0 \langle \Sigma \rangle$ is the tree joining 0 and i .
- If $\langle \Sigma \rangle = (\langle \Sigma_1 \rangle \boxplus \langle \Sigma_2 \rangle)$ then $T_0 \langle \Sigma \rangle$ is the tree obtained by identifying the vertices 0 in $T_0 \langle \Sigma_1 \rangle$ and $T_0 \langle \Sigma_2 \rangle$ to a single vertex x (corresponding to the instance of \boxplus operating on $\langle \Sigma_1 \rangle$ and $\langle \Sigma_2 \rangle$), and joining x to a new root 0.

It is clear that two parenthesizations of Σ are equivalent if and only if their associated rooted abstract Steiner trees are equal.

The solution to Schröder's third problem, i.e., the number a_k of equivalence classes of parenthesizations with support $[k]$, is the product of the first $k - 1$ odd numbers:

$$a_k = \prod_{i=1}^{k-1} (2i - 1).$$

This is seen as follows. Since a_k equals the number of rooted abstract Steiner trees on $[k]$, i.e., trees with $k+1$ leaves $\{0, 1, \dots, k\}$ and $k-1$ internal vertices of degree 3, we have $a_1 = 1$ and $a_{k+1} = (2k-1)a_k$, since we may subdivide any of the $2k-1$ edges of such a tree on $[k]$ and join the new vertex to $k+1$, to obtain such a tree on $[k+1]$.

We remark that the k th Catalan number can be derived from a_k . Since there are $k-1$ \boxplus 's in a parenthesization of $\{A_i : i \in [k]\}$, there are 2^{k-1} parenthesizations in an equivalence class. This gives $2^{k-1}a_k$ parenthesizations (in the usual sense) of the ordered expression $A_{\pi(1)} \boxplus \dots \boxplus A_{\pi(k)}$, where π is a permutation of $[k]$, taken over all permutations π . Therefore, the number of parenthesizations of $A_1 \boxplus \dots \boxplus A_k$ is $2^{k-1}a_k/k!$, which equals the k th Catalan number. This derivation of the Catalan numbers is essentially the same as the classical combinatorial derivation of Rodrigues [32].

We define an *abstract Steiner tree* on I to be a tree with set of leaves I , and with $|I| - 2$ internal vertices of degree 3. The abstract Steiner tree associated to a rooted abstract Steiner tree on I is obtained by contracting the root 0 and its incident edge. The abstract Steiner tree of a parenthesization $\langle \Sigma \rangle$ obtained in this way from $T_0\langle \Sigma \rangle$ is denoted by $T\langle \Sigma \rangle$. The number of abstract Steiner trees on $[k]$ clearly equals a_{k-1} .

We call two parenthesizations *weakly equivalent* if they can be transformed into each other using the commutative law on any subexpression or the associative law on the whole expression, i.e., a bracketed expression

$$((\langle \Sigma_1 \rangle \boxplus \langle \Sigma_2 \rangle) \boxplus \langle \Sigma_3 \rangle)$$

may be transformed into

$$(\langle \Sigma_1 \rangle \boxplus (\langle \Sigma_2 \boxplus \Sigma_3 \rangle)),$$

and vice versa. It is again clear that two parenthesizations Σ and Σ' are weakly equivalent if and only if their abstract Steiner trees are equal: $T\langle \Sigma \rangle = T\langle \Sigma' \rangle$. It follows that the number of weak equivalence classes of parenthesizations on $[k]$ equals a_{k-1} . The operation \boxplus has the following weak associativity property.

PROPOSITION 14. *Let $\langle \Sigma \rangle_1$ and $\langle \Sigma \rangle_2$ be two weakly equivalent parenthesizations of $\Sigma = \{A_i : i \in I\}$, where each A_i is a closed convex subset of $B(\mathbb{M}^*)$. Then $o \in \langle \Sigma \rangle_1$ if and only if $o \in \langle \Sigma \rangle_2$.*

PROOF. By the definition of weak equivalence, it is sufficient to show for any three $A_i \subseteq B(\mathbb{M}^*)$, $i = 1, 2, 3$, that $o \in A_1 \boxplus (A_2 \boxplus A_3)$ if and only if $o \in (A_1 \boxplus A_2) \boxplus A_3$. However, since $A_i \subseteq B(\mathbb{M}^*)$, both statements are equivalent to $o \in A_1 + A_2 + A_3$. \square

4.2. The characterization.

THEOREM 15 (Nodes). *Let $N = \{\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_k\}$ be a set of points in a Minkowski space \mathbb{M}^n . Then the star joining \mathbf{p}_0 to each \mathbf{p}_i , $i \in [k]$, is an SMT of N if and only if $\langle \Sigma \rangle \neq \emptyset$ for each parenthesization $\langle \Sigma \rangle$ of $\Sigma = \{\partial(\mathbf{p}_i - \mathbf{p}_0) : i \in [k]\}$.*

PROOF. \Rightarrow : Consider any parenthesization $\langle \Sigma \rangle$ of Σ and its associated rooted abstract Steiner tree $T_0\langle \Sigma \rangle$. We now turn this tree into a Steiner tree in \mathbb{M}^n . Associate leaf i with \mathbf{p}_i for each $i \in \{0\} \cup [k]$, and associate each of the $k-1$ internal vertices with a variable point $\mathbf{x}_i \in \mathbb{M}^n$, $i \in [k-1]$. Denote this Steiner

tree by $T_0\langle\Sigma\rangle(\mathbf{x}_1, \dots, \mathbf{x}_{k-1})$. Its length is

$$L(\mathbf{x}_1, \dots, \mathbf{x}_{k-1}) := \ell(T_0\langle\Sigma\rangle(\mathbf{x}_1, \dots, \mathbf{x}_{k-1})).$$

The Steiner points \mathbf{x}_i in $T_0\langle\Sigma\rangle(\mathbf{x}_1, \dots, \mathbf{x}_{k-1})$ may coincide — this results in the tree being in fact a contraction of $T_0\langle\Sigma\rangle$. Note that $L : \mathbb{M}^n \times \dots \times \mathbb{M}^n \rightarrow \mathbb{R}$ is a convex function since

$$L(\mathbf{x}_1, \dots, \mathbf{x}_{k-1}) = \sum_{e \in E(T_0\langle\Sigma\rangle)} \rho_e(\mathbf{x}_1, \dots, \mathbf{x}_{k-1}),$$

where $\rho_e(\mathbf{x}_1, \dots, \mathbf{x}_{k-1}) = \|\mathbf{a} - \mathbf{b}\|$, with \mathbf{a} and \mathbf{b} the two points in $N \cup \{\mathbf{x}_1, \dots, \mathbf{x}_{k-1}\}$ associated to the vertices incident to e . Each ρ_e depends on only one or two of the variables \mathbf{x}_i . Since $T_0\langle\Sigma\rangle(\mathbf{p}_0, \dots, \mathbf{p}_0)$ is the star joining \mathbf{p}_0 to all \mathbf{p}_i , $i \in [k]$, which is an SMT by assumption, L attains its minimum at $(\mathbf{p}_0, \dots, \mathbf{p}_0)$. Thus

$$o \in \partial L(\mathbf{p}_0, \dots, \mathbf{p}_0) = \sum_{e \in E(T_0\langle\Sigma\rangle)} \partial \rho_e(\mathbf{p}_0, \dots, \mathbf{p}_0) \quad (5)$$

by Lemma 10. By Lemmas 11, 12 and 13, if $e = x_i x_j$, then

$$\partial \rho_e(\mathbf{p}_0, \dots, \mathbf{p}_0) = \{(o, \dots, o, \underbrace{\varphi}_{\text{pos. } i}, o, \dots, o, \underbrace{-\varphi}_{\text{pos. } j}, o, \dots, o) : \varphi \in B^*\}$$

while if $e = p_i x_j$, then

$$\partial \rho_e(\mathbf{p}_0, \dots, \mathbf{p}_0) = \{(o, \dots, o, \underbrace{\varphi}_{\text{pos. } j}, o, \dots, o) : \varphi \in \partial(\mathbf{p}_i - \mathbf{p}_0)\}.$$

By considering each coordinate $i \in [k-1]$ of (5) we obtain a functional $\varphi_e \in (M^n)^*$ for each edge $e \in E(T_0\langle\Sigma\rangle)$ such that

- $\varphi_e \in \begin{cases} B^* & \text{if } e = \mathbf{x}_i \mathbf{x}_j \text{ or } e = \mathbf{p}_0 \mathbf{x}_i, \\ \partial(\mathbf{p}_i - \mathbf{p}_0) & \text{if } e = \mathbf{p}_i \mathbf{x}_j, i \neq 0, \end{cases}$
- and for each Steiner point \mathbf{x}_i , $\varphi_e = \varphi_f + \varphi_g$, where e is the incoming edge and f, g the two outgoing edges of \mathbf{x}_i , when the tree is directed away from the root \mathbf{p}_0 .

By induction on the definition of $T_0\langle\Sigma\rangle$ (≡ induction on subexpressions of $\langle\Sigma\rangle$) we obtain that $\varphi_e \in \langle\Sigma\rangle$, where $e = \mathbf{p}_0 \mathbf{x}_i$ is the root edge. This gives $\langle\Sigma\rangle \neq \emptyset$.

\Leftarrow : Consider any Steiner tree in \mathbb{M}^n on $\{\mathbf{p}_0, \dots, \mathbf{p}_k\}$. By subdividing points if necessary, we obtain a tree with leaves $\{\mathbf{p}_0, \dots, \mathbf{p}_k\}$ and with $k-1$ Steiner points \mathbf{x}_i of degree 3, some of them possibly coinciding with each other or with the \mathbf{p}_i . This tree is the rooted abstract Steiner tree of some parenthesization $\langle\Sigma\rangle$ of Σ . As in the “ \Rightarrow ”-argument, we obtain that $\langle\Sigma\rangle \neq \emptyset$ implies that $o \in \partial L(\mathbf{p}_0, \dots, \mathbf{p}_0)$, i.e., L attains its minimum at $(\mathbf{p}_0, \dots, \mathbf{p}_0)$, which is when the tree is $T_0\langle\Sigma\rangle(\mathbf{p}_0, \dots, \mathbf{p}_0)$, the star joining \mathbf{p}_0 to the other \mathbf{p}_i . \square

THEOREM 16 (Steiner points). *Let $N = \{\mathbf{p}_1, \dots, \mathbf{p}_k\}$ be a set of points in a Minkowski space \mathbb{M}^n . Then the star joining \mathbf{p}_0 to each \mathbf{p}_i , $i \in [k]$, is an SMT of N if and only if $o \in \langle\Sigma\rangle$ for each parenthesization $\langle\Sigma\rangle$ of $\Sigma = \{\partial(\mathbf{p}_i - \mathbf{p}_0) : i \in [k]\}$, if and only if $o \in \langle\Sigma'\rangle + \partial(\mathbf{p}_k - \mathbf{p}_0)$ for each parenthesization $\langle\Sigma'\rangle$ of $\Sigma' = \{\partial(\mathbf{p}_i - \mathbf{p}_0) : i \in [k-1]\}$.*

PROOF. Since any parenthesization $\langle \Sigma \rangle$ is weakly equivalent to $\langle \Sigma' \rangle \boxplus \partial(\mathbf{p}_k - \mathbf{p}_0)$ for some parenthesization of $\langle \Sigma' \rangle$, Proposition 14 gives the equivalence between the two conditions. We next show that the first condition is necessary and sufficient.

\Rightarrow : This is similar to the proof of Theorem 15. Consider any parenthesization $\langle \Sigma \rangle$ of Σ . Turn its associated rooted abstract Steiner tree $T_0\langle \Sigma \rangle$ into an abstract Steiner tree $T\langle \Sigma \rangle$ by contracting the root. We turn $T\langle \Sigma \rangle$ into a directed graph as follows. Denote the (new) edge into which the root was contracted by \bar{e} , and give it both directions, denoting the two directed edges by e^+ and e^- . Give all other edges of $T\langle \Sigma \rangle$ a single direction away from \bar{e} . This tree becomes a Steiner tree in \mathbb{M}^n as follows. Associate leaf i with \mathbf{p}_i for each $i \in [k]$, and associate each of the $k - 2$ internal vertices with a variable point $\mathbf{x}_i \in \mathbb{M}^n$, $i \in [k - 2]$. Denote this Steiner tree by $T\langle \Sigma \rangle(\mathbf{x}_1, \dots, \mathbf{x}_{k-2})$. Its length is the convex function

$$L(\mathbf{x}_1, \dots, \mathbf{x}_{k-2}) := \ell(T\langle \Sigma \rangle(\mathbf{x}_1, \dots, \mathbf{x}_{k-2})).$$

Again note that the Steiner points \mathbf{x}_i in $T\langle \Sigma \rangle(\mathbf{x}_1, \dots, \mathbf{x}_{k-2})$ may coincide, and then the tree is a contraction of $T\langle \Sigma \rangle$. Again, L attains its minimum at $(\mathbf{p}_0, \dots, \mathbf{p}_0)$. Calculating the subdifferential coordinate-wise, we obtain a functional $\varphi_{\bar{e}} \in (\mathbb{M}^n)^*$ for each directed edge \bar{e} of $E(T\langle \Sigma \rangle)$ such that

- $\varphi_{\bar{e}} \in \begin{cases} B^* & \text{if } \bar{e} \text{ is incident with two Steiner points,} \\ \partial(\mathbf{p}_i - \mathbf{p}_0) & \text{if } \bar{e} \text{ is incident with } \mathbf{p}_i, i \neq 0, \end{cases}$
- $\varphi_{e^+} = -\varphi_{e^-}$, and
- for each Steiner point \mathbf{x}_i , $\varphi_{\bar{e}} = \varphi_{\bar{f}} + \varphi_{\bar{g}}$, where \bar{e} is the incoming edge and \bar{f}, \bar{g} the two outgoing edges of \mathbf{x}_i , with the convention that we ignore the outgoing e^+ or e^- if \mathbf{x}_i is incident with \bar{e} .

Write $\langle \Sigma \rangle = \langle \Sigma^+ \rangle \boxplus \langle \Sigma^- \rangle$, where e^\pm points to the subtree associated with $\langle \Sigma^\pm \rangle$. Let I^\pm be the support of $\langle \Sigma^\pm \rangle$, and for each $i \in [k]$ let $\varphi_i = \varphi_{\bar{e}}$, where \bar{e} is incident with \mathbf{x}_i . Again by induction on subexpressions we obtain that $\varphi_{e^\pm} = \sum_{i \in I^\pm} \varphi_i \in \langle \Sigma^\pm \rangle$. From $\varphi_{e^+} = -\varphi_{e^-}$ it follows that $o \in \langle \Sigma \rangle$.

\Leftarrow : Similar to the corresponding direction in the proof of Theorem 15. \square

By the discussion in Section 4.1, when applying Theorem 15 (Theorem 16) there are a_k (a_{k-1} , respectively) parenthesizations to consider.

Note that in the “ \Leftarrow ”-directions of the above proofs we did not need the parts of Lemmas 9 to 13 depending on the separation theorem, i.e., Lemma 9 and the “ \subseteq ”-part of Lemma 10. These directions of Theorems 15 and 16 are used to obtain lower bounds for $d(\mathbb{M}^n)$ and $s(\mathbb{M}^n)$. On the other hand, to obtain upper bounds we need the “ \Rightarrow ”-directions, where the separation theorem is needed.

Recall that if the norm is differentiable, then $\partial\|\mathbf{x}\|$ is a singleton whenever $\mathbf{x} \neq \mathbf{o}$. This drastically simplifies the conditions in Theorems 15 and 16 and we regain the following two results.

COROLLARY 17 (Lawlor and Morgan [21]). *Let $N = \{\mathbf{p}_1, \dots, \mathbf{p}_k\}$ be a set of points in a Minkowski space \mathbb{M}^n with differentiable norm. Let φ_i be the norming functional of $\mathbf{p}_i - \mathbf{p}_0$, $i \in [k]$. Then the star joining \mathbf{p}_0 to each \mathbf{p}_i , $i \in [k]$, is an SMT of N if and only if*

$$\sum_{i=1}^k \varphi_i = o$$

and for each subset $I \subseteq [k]$,

$$\left\| \sum_{i \in I} \varphi \right\|^* \leq 1.$$

COROLLARY 18 ([36]). *Let $N = \{\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_k\}$ be a set of points in a Minkowski space \mathbb{M}^n with differentiable norm. Let φ_i be the norming functional of $\mathbf{p}_i - \mathbf{p}_0$, $i = 1, \dots, k$. Then the star joining \mathbf{p}_0 to each \mathbf{p}_i , $i = 1, \dots, k$, is an SMT of N if and only if for each subset $I \subseteq [k]$,*

$$\left\| \sum_{i \in I} \varphi \right\|^* \leq 1.$$

5. Using extremal set theory to prove Theorem 3

In order to determine $d(\mathbb{M}_Z^n)$ we apply Theorem 15. It follows from the discussion in Section 2.4 that norming functionals are described as follows.

LEMMA 19. *For any non-zero $\mathbf{x} \in \mathbb{M}_Z^n$, we have $\partial\mathbf{x} = F'(X)$, where $X \in \mathcal{Z}_n$ is the unique signed set such that $F(X)$ is the face of Z_n which has $\frac{1}{\|\mathbf{x}\|_Z} \mathbf{x}$ in its relative interior.*

We thus have to determine for which families $\Sigma = \{X_i : i \in [k]\} \subseteq \mathcal{Z}_n$ all parenthesizations $\langle F'(\Sigma) \rangle$ of $F'(\Sigma) := \{F'(X_i) : i \in [k]\}$ are non-empty. To this end, we define a commutative, non-associative binary operation \square on $\overline{\mathcal{Z}_n}$ as follows:

- $X \square \mathbf{1} = \mathbf{1} \square X = X$ for all $X \in \overline{\mathcal{Z}_n}$, and
- for all $X, Y \neq \mathbf{1}$,

$$X \square Y := \begin{cases} (X^+, Y^-) & \text{if } X^+ \cap Y^- = \emptyset \text{ and } X^- \cap Y^+ \neq \emptyset, \\ (Y^+, X^-) & \text{if } X^+ \cap Y^- \neq \emptyset \text{ and } X^- \cap Y^+ = \emptyset, \\ \mathbf{1} & \text{if } X^+ \cap Y^- \neq \emptyset \text{ and } X^- \cap Y^+ \neq \emptyset, \\ \emptyset & \text{if } X^+ \cap Y^- = \emptyset \text{ and } X^- \cap Y^+ = \emptyset. \end{cases}$$

Extending F' to $\overline{\mathcal{Z}_n}$ by defining $F'(\mathbf{1}) = \{o\} \subset \mathbb{M}_C^n$, we obtain the following.

LEMMA 20. *For all $X, Y \in \overline{\mathcal{Z}_n}$,*

$$F'(X \square Y) \subseteq F'(X) \boxplus F'(Y).$$

Furthermore, if at least one of the conditions $X = \mathbf{1}$, $Y = \mathbf{1}$, $X^+ \cap Y^- = \emptyset$ or $X^- \cap Y^+ = \emptyset$ holds, then

$$F'(X \square Y) = F'(X) \boxplus F'(Y).$$

PROOF. We assume that $X, Y \neq \mathbf{1}$, otherwise the lemma is trivial.

Consider the first relation. Since this is trivial if $X \square Y = \emptyset$, we assume that X and Y are not conformal, say that $X^+ \cap Y^- \neq \emptyset$. There are now two cases, depending on whether $X^- \cap Y^+$ is empty or not.

- If $X^- \cap Y^+ = \emptyset$, we have to show that $F'(Y^+, X^-) \subseteq F'(X) \boxplus F'(Y)$. Choose any $i \in X^+ \cap Y^-$. Then for any $\varphi \in F'(Y^+, X^-)$ we have $\varphi(\mathbf{e}_i) = 0$. By setting $\psi = \varphi^- + \frac{1}{2}\epsilon_i \in F'(X)$ and $\chi = \varphi^+ - \frac{1}{2}\epsilon_i \in F'(Y)$, we obtain $\varphi = \psi + \chi \in F'(X) \boxplus F'(Y)$.
- If $X^- \cap Y^+ \neq \emptyset$, we have to show that $F'(\mathbf{1}) \subseteq F'(X) \boxplus F'(Y)$. Choose any $i \in X^+ \cap Y^-$ and $j \in X^- \cap Y^+$. Then $\frac{1}{2}\epsilon_i - \frac{1}{2}\epsilon_j \in F'(X)$ and $-\frac{1}{2}\epsilon_i + \frac{1}{2}\epsilon_j \in F'(Y)$, giving $o \in F'(X) \boxplus F'(Y)$.

This establishes the inclusion.

Now consider the equality. Without loss assume $X^+ \cap Y^- = \emptyset$. Let $\varphi \in F'(X)$ and $\psi \in F'(Y)$ be such that $\|\varphi + \psi\|_1 \leq 1$. Then for each $i \in X^+$ we have $0 \leq \psi(\mathbf{e}_i) \leq \varphi(\mathbf{e}_i) + \psi(\mathbf{e}_i)$ and for each $i \in Y^-$ we have $0 \geq \varphi(\mathbf{e}_i) \geq \varphi(\mathbf{e}_i) + \psi(\mathbf{e}_i)$. In particular, $X^+ \cap \text{supp}^-(\varphi + \psi) = \emptyset$ and $Y^- \cap \text{supp}^+(\varphi + \psi) = \emptyset$, and

$$\begin{aligned} 1 &\geq \|\varphi + \psi\|_1 \\ &= \sum_{i \in \text{supp}^+(\varphi + \psi)} (\varphi(\mathbf{e}_i) + \psi(\mathbf{e}_i)) + \sum_{i \in \text{supp}^-(\varphi + \psi)} -(\varphi(\mathbf{e}_i) + \psi(\mathbf{e}_i)) \\ &\geq \sum_{i \in X^+} (\varphi(\mathbf{e}_i) + \psi(\mathbf{e}_i)) + \sum_{i \in Y^-} -(\varphi(\mathbf{e}_i) + \psi(\mathbf{e}_i)) \\ &= \frac{1}{2} + \sum_{i \in X^+} \psi(\mathbf{e}_i) + \frac{1}{2} + \sum_{i \in Y^-} -\varphi(\mathbf{e}_i) \\ &\geq 1. \end{aligned}$$

Thus equality holds everywhere, giving $\|\varphi + \psi\|_1 = 1$, $\text{supp}^+(\varphi + \psi) \subseteq X^+$, and $\text{supp}^-(\varphi + \psi) \subseteq Y^-$. Therefore, $\varphi + \psi \in F'(X^+, Y^-)$. If $X^- \cap Y^+ \neq \emptyset$, then $(X^+, Y^-) = X \square Y$, and we are done. Otherwise, $X^- \cap Y^+ = \emptyset$, giving that X and Y are conformal. Then $F'(X \square Y) = \emptyset$, and also $F'(X) \boxplus F'(Y) = \emptyset$, since $F'(X)$ and $F'(Y)$ belong to the same facet of Z_n^* . \square

THEOREM 21. *Let $\Sigma = \{X_i : i \in [k]\} \subseteq \mathcal{Z}_n \setminus \{\emptyset\}$. Then the following are equivalent.*

- (a) *Some parenthesization of $F'(\Sigma)$ is empty.*
- (b) *Some parenthesization of Σ (with operation \square) equals \emptyset .*
- (c) *There exist indices $a, b, c, d \in [k]$ with $\{a, b\} \cap \{c, d\} = \emptyset$ such that $(X_a^+ \cup X_c^+) \cap (X_b^- \cup X_d^-) = \emptyset$.*

PROOF. The implication (a) \implies (b) follows from Lemma 20, since $F'(X) = \emptyset$ implies $X = \emptyset$.

(b) \implies (c): Let $\langle \Sigma \rangle = \emptyset$. By the definition of \square , for some subexpression $\langle \Sigma_1 \rangle \square \langle \Sigma_2 \rangle$ of $\langle \Sigma \rangle$ we have $\langle \Sigma_1 \rangle \neq \emptyset \neq \langle \Sigma_2 \rangle$ but $\langle \Sigma_1 \rangle \square \langle \Sigma_2 \rangle = \emptyset$. If both Σ_1 and Σ_2 are singletons, say $\Sigma_1 = \{X_i\}$ and $\Sigma_2 = \{X_j\}$, then X_i and X_j are conformal, i.e., $X_i^+ \cap X_j^- = \emptyset = X_j^+ \cap X_i^-$, so we may take $a = b = i$ and $c = d = j$.

If Σ_1 is a singleton $\{X_a\}$ but Σ_2 is not, then by induction on the definition of \square there exist $X_c, X_d \in \Sigma_2$ such that $\langle \Sigma_2 \rangle = (X_c^+, X_d^-)$ is conformal with X_a , i.e., $X_a^+ \cap X_d^- = \emptyset = X_c^+ \cap X_a^-$. Thus we may set $b = a$. A similar argument takes care of the case where Σ_2 is a singleton.

If Σ_1 and Σ_2 are both not singletons, then again by induction on the definition of \square there exist $X_a, X_b \in \Sigma_1$ and $X_c, X_d \in \Sigma_2$ such that $\langle \Sigma_1 \rangle = (X_a^+, X_b^-)$ and $\langle \Sigma_2 \rangle = (X_c^+, X_d^-)$ are conformal, giving (c).

(c) \implies (a): It is sufficient to find a parenthesization of a subfamily of $F'(\Sigma)$ that equals the empty set. First consider the case where a, b, c, d are distinct indices. Without loss assume that $F'(X_a) \boxplus F'(X_b) \neq \emptyset \neq F'(X_c) \boxplus F'(X_d)$. By Lemma 20 we have $F'(X_a) \boxplus F'(X_b) = F'(X_a \square X_b)$, hence $X_a \square X_b \neq \emptyset$, and $X_a \square X_b = (X_a^+, X_b^-)$. Similarly, $F'(X_c) \boxplus F'(X_d) = F'(X_c^+, X_d^-)$, and again by

Lemma 20,

$$\begin{aligned}
& (F'(X_a) \boxplus F'(X_b)) \boxplus (F'(X_c) \boxplus F'(X_d)) \\
&= F'(X_a^+, X_b^-) \boxplus F'(X_c^+, X_d^-) \\
&= F'(\emptyset) = \emptyset.
\end{aligned}$$

Similarly,

- if $a = b$ and $c = d$, then $F'(X_a) \boxplus F'(X_c) = \emptyset$,
- if $a = b$ and $c \neq d$, then $F'(X_a) \boxplus (F'(X_c) \boxplus F'(X_d)) = \emptyset$, and
- if $a \neq b$ and $c = d$, then $(F'(X_a) \boxplus F'(X_b)) \boxplus F'(X_c) = \emptyset$. \square

The above theorem together with Theorem 15 now gives the following.

COROLLARY 22. *Let $P = \{\mathbf{p}_i : i \in [k]\}$ be a family of points in $\mathbb{M}_Z^n \setminus \{\mathbf{o}\}$, with $\|\mathbf{p}_i\|_Z^{-1} \mathbf{p}_i$ in the relative interior of face $F(X_i)$ of Z_n . Then the star connecting P to \mathbf{o} is an SMT of $P \cup \{\mathbf{o}\}$ if and only if there do not exist indices $a, b, c, d \in [k]$ with $\{a, b\} \cap \{c, d\} = \emptyset$ and $(X_a^+ \cup X_c^+) \cap (X_b^- \cup X_d^-) = \emptyset$.*

Consequently, the star connecting P to \mathbf{o} is an SMT of $P \cup \{\mathbf{o}\}$ if and only if for any $I \subseteq [k]$, $2 \leq |I| \leq 3$, the star connecting $\{\mathbf{p}_i : i \in I\}$ to \mathbf{o} is an SMT of $\{\mathbf{p}_i : i \in I\} \cup \{\mathbf{o}\}$.

This also implies Corollary 4. The problem of determining $d(\mathbb{M}_Z^n)$ has now been reduced to a problem in extremal finite set theory.

THEOREM 23. *Let $\{X_i : i \in [k]\}$ be a family of signed sets from \mathcal{Z}_n such that there do not exist indices $a, b, c, d \in [k]$ with $\{a, b\} \cap \{c, d\} = \emptyset$ and $(X_a^+ \cup X_c^+) \cap (X_b^- \cup X_d^-) = \emptyset$. Then $k \leq \binom{n+2}{\lfloor (n+2)/2 \rfloor}$ with equality if and only if all $X_i^0 = \emptyset$, and either all $|X_i^+| \in \{\lfloor (n+1)/2 \rfloor, \lfloor (n+1)/2 \rfloor + 1\}$, or all $|X_i^+| \in \{\lceil (n+1)/2 \rceil - 1, \lceil (n+1)/2 \rceil\}$.*

PROOF. It is easily seen that the hypothesis is equivalent to the following statement: For all families of sets $\{Y_i : i \in [k]\} \subseteq \mathcal{P}[n+1] \setminus \{\emptyset, [n+1]\}$ such that $X_i^+ \subseteq Y_i \subseteq [n+1] \setminus X_i^-$,

$$\begin{aligned}
& \text{there do not exist indices } a, b, c, d \in [k] \\
& \text{with } \{a, b\} \cap \{c, d\} = \emptyset \text{ and } Y_a \cup Y_c \subseteq Y_b \cap Y_d.
\end{aligned} \tag{6}$$

Property (6) is in turn equivalent to the following three conditions:

$$\text{all } Y_i \text{ are distinct sets,} \tag{7}$$

$$\text{there do not exist distinct } a, b, c \in [k] \text{ with } Y_a \subseteq Y_b \subseteq Y_c, \text{ and} \tag{8}$$

$$\text{there do not exist distinct } a, b, c, d \in [k] \text{ with } Y_a \cup Y_b \subseteq Y_c \cap Y_d. \tag{9}$$

By a well-known generalisation of Sperner's theorem due to Erdős [13, 2], conditions (7) and (8) on their own already give the sharp upper bound $k \leq \binom{n+2}{\lfloor (n+2)/2 \rfloor}$ with equality exactly when $\{Y_i : i \in [k]\}$ consists of two of the largest levels in $\mathcal{P}[n+1]$. Because of this rigidity, it easily follows that in the case of equality all $X_i^0 = \emptyset$, finishing the proof. \square

Theorem 3 is now proved. We remark that conditions (7) and (9) on their own give the same upper bound, but with an extra case of equality when $n+1 = 4$; see [11].

6. A geometric formulation of Sperner's theorem

The next lemma follows from an observation of Moore [15].

LEMMA 24. *If the unit ball of \mathbb{M}^n contains k points on its boundary such that the distance between any two equals 2, then $s(\mathbb{M}^n) \geq k$.*

From this lemma together with Lemma 2 it immediately follows that $s(\ell_\infty^n) = d(\ell_\infty^n) = 2^n$ and $s(\ell_1^n) = d(\ell_1^n) = 2n$. We apply this lemma to \mathbb{M}_Z^n . Recall that a vertex of Z_n equals some $F(X)$, where $X^0 = \emptyset$ and $X^\pm \neq \emptyset$. Thus X can be identified with the set $X^+ \subseteq [n+1]$, not equal to \emptyset or $[n+1]$. We then have for any distinct $X^+, Y^+ \in \mathcal{P}[n+1] \setminus \{\emptyset, [n+1]\}$ that

$$\|F(X^+) - F(Y^+)\|_Z = \begin{cases} 1 & \text{if } X^+ \subset Y^+ \text{ or } Y^+ \subset X^+, \\ 2 & \text{if } X^+ \not\subseteq Y^+ \text{ and } Y^+ \not\subseteq X^+. \end{cases}$$

It follows that a set $\{F(X_i) : i \in I\}$ of vertices of Z_n are all at pairwise distance 2 if and only if $\{X_i^+ : i \in I\}$ is an antichain. It follows from Lemma 24 that $s(\mathbb{M}_Z^n) \geq \binom{n+1}{\lfloor (n+1)/2 \rfloor}$. This establishes Theorem 6.

By Sperner's theorem [2] we cannot do better: The largest set of vertices at pairwise distance 2 has size $\binom{n+1}{\lfloor (n+1)/2 \rfloor}$. We still cannot do better even if we consider arbitrary points on the boundary of Z_n . For any boundary point \mathbf{p} of Z_n its support $X_{\mathbf{p}} := \text{supp}(\mathbf{p})$ describes the unique face $F(X_{\mathbf{p}})$ which contains \mathbf{p} in its relative interior. Two boundary points \mathbf{p} and \mathbf{q} are at distance 2 if and only if there exist parallel supporting hyperplanes at \mathbf{p} and \mathbf{q} with Z_n in between. This in turn is equivalent to \mathbf{p} and $-\mathbf{q}$ being contained in the same facet $F(X)$ of Z_n . Note that $F(X)$ is a facet of Z_n if and only if $|X^\pm| = 1$. Therefore, $\|\mathbf{p} - \mathbf{q}\|_Z = 2$ if and only if $X_{\mathbf{p}}^+ \cap X_{\mathbf{q}}^- \neq \emptyset$ and $X_{\mathbf{p}}^- \cap X_{\mathbf{q}}^+ \neq \emptyset$. Now let $P = \{\mathbf{p}_i : i \in I\}$ be a set of boundary points of Z_n at pairwise distance 2, let $X_i = \text{supp}(\mathbf{p}_i)$, and choose $Y_i \in \mathcal{P}[n+1] \setminus \{\emptyset, [n+1]\}$ such that $X_i^+ \subseteq Y_i \subseteq [n+1] \setminus X_i^-$. Then it follows that the Y_i are all distinct sets, and form an antichain. As before we have by Sperner's theorem that $|P| \leq \binom{n+1}{\lfloor (n+1)/2 \rfloor}$, with equality if and only if all \mathbf{p}_i are vertices and all X_i^+ have the same cardinality, either $\lfloor (n+1)/2 \rfloor$ or $\lceil (n+1)/2 \rceil$. We have shown the following.

PROPOSITION 25. *The largest number of unit vectors in \mathbb{M}_Z^n at pairwise distance 2 is $\binom{n+1}{\lfloor (n+1)/2 \rfloor}$. \square*

An even more general result would be the following.

CONJECTURE 26. *The largest size of an equilateral set in \mathbb{M}_Z^n equals $\binom{n+1}{\lfloor (n+1)/2 \rfloor}$.*

This conjecture is known to hold for $n = 3$ [34] (and is easy for $n = 2$).

7. Perturbations of ℓ_1^n

The following lemma is standard.

LEMMA 27. *Let $\|\cdot\|_a$ and $\|\cdot\|_b$ be any norms on \mathbb{R}^n . Then the dual of $\mathbb{M}^n = (\mathbb{R}^n, \|\cdot\|_a + \|\cdot\|_b)$ is isometrically isomorphic to $((\mathbb{R}^n)^*, B_a^* + B_b^*)$, where B_a^* and B_b^* are the unit balls of the dual norms $\|\cdot\|_a^*$ and $\|\cdot\|_b^*$.*

PROOF OF THEOREM 7. Let $\|\cdot\| = \|\cdot\|_1 + \lambda\|\cdot\|_2$. By Lemma 27 we may identify the dual unit ball B^* with $\sum_{i=1}^n [-\epsilon_i, \epsilon_i] + \lambda B_2^*$. Recall that each $\partial\|\mathbf{x}\|$, $\mathbf{x} \neq \mathbf{o}$, is an exposed face of B^* . Any exposed face of B^* equals a proper non-empty face F of the cube $\sum_{i=1}^n [-\epsilon_i, \epsilon_i]$ translated by a functional φ with $\|\varphi\|_2^* = \lambda$, such that F and φ have the same sign in the following sense: if $F = \sum_{i \in X^+} \epsilon_i - \sum_{i \in X^-} \epsilon_i + \sum_{i \in X^0} [-\epsilon_i, \epsilon_i]$ for some signed set $X \neq \emptyset$, then $X = \text{supp}(\varphi)$.

We first show $d(\mathbb{R}^n, \|\cdot\|) \leq 2n$ if $\lambda \leq 1$. Let $\mathbf{p}_1, \dots, \mathbf{p}_k \in \mathbb{M}^n$, $\mathbf{p}_i \neq \mathbf{o}$, and suppose that the star joining \mathbf{o} to all \mathbf{p}_i is an SMT of $\{\mathbf{o}\} \cup \{\mathbf{p}_i : i \in [k]\}$. Then by Theorem 15 all restricted Minkowski sums of the $\partial\|\mathbf{p}_i\|$ must be non-empty. Let $\partial\|\mathbf{p}_i\| = F_i + \varphi_i$ as above, with X_i the corresponding signed set. Suppose $k > 2n$. Since all $X_i \neq \emptyset$, it follows from the pigeon-hole principle that for some two indices i, j , X_i and X_j have a common element of the same sign, say $1 \in X_1^+ \cap X_2^+$. Then, since F_1 and F_2 are both contained in the hyperplane $\{\chi : \chi(\mathbf{e}_1) = 1\}$, it follows that $\partial\|\mathbf{p}_1\| + \partial\|\mathbf{p}_2\|$ is contained in the hyperplane $\{\chi : \chi(\mathbf{e}_1) = 2 + \varphi_1(\mathbf{e}_1) + \varphi_2(\mathbf{e}_1)\}$. However, B^* is contained in the slab bounded by $\{\chi : \chi(\mathbf{e}_1) = 1 + \lambda\}$. Since $2 + \varphi_1(\mathbf{e}_1) + \varphi_2(\mathbf{e}_1) > 2 \geq 1 + \lambda$, we obtain $\partial\|\mathbf{p}_1\| \boxplus \partial\|\mathbf{p}_2\| = \emptyset$, a contradiction. Therefore, $k \leq 2n$.

We now prove that $s(\mathbb{R}^n, \|\cdot\|) \geq 2n$ for $\lambda \leq \sqrt{n}/(\sqrt{n} - 1)$ by showing that the star joining \mathbf{o} to all $\pm \mathbf{e}_i$, $i \in [n]$, is an SMT of $\{\pm \mathbf{e}_i : i \in [n]\}$. This is trivial for $n = 1$, so we assume from now on that $n \geq 2$. We have

$$\partial\|\mathbf{e}_i\| = E_i := \sum_{\substack{j=1 \\ j \neq i}}^n [-\epsilon_j, \epsilon_j] + (1 + \lambda)\epsilon_i.$$

By Theorem 16 it is sufficient to prove that for each parenthesization $\langle \Sigma \rangle$ of

$$\Sigma := \{\pm E_i : i \in [n-1]\} \cup \{E_n\}$$

we have $\mathbf{o} \in \langle \Sigma \rangle - E_n$, or equivalently, $E_n \cap \langle \Sigma \rangle \neq \emptyset$.

Write

$$\langle \Sigma \rangle = \langle \Sigma_1 \rangle \boxplus \langle \Sigma_2 \rangle = (\langle \Sigma_1 \rangle + \langle \Sigma_2 \rangle) \cap B^*$$

where $\Sigma_1 \cup \Sigma_2 = \Sigma$, $\Sigma_1 \cap \Sigma_2 = \emptyset$, and $E_n \in \Sigma_1$. Since $E_n \subseteq B^*$, we only have to prove that $E_n \cap (\langle \Sigma_1 \rangle + \langle \Sigma_2 \rangle) \neq \emptyset$. We now replace the operation \boxplus in $\langle \Sigma_i \rangle$ by \boxplus' , where

$$C \boxplus' D := (C + D) \cap [-1 - \lambda/\sqrt{n}, 1 + \lambda/\sqrt{n}]^n.$$

Then $C \boxplus' D \subseteq C \boxplus D$, since $[-1 - \lambda/\sqrt{n}, 1 + \lambda/\sqrt{n}]^n \subseteq B^*$. Denoting the parenthesizations with respect to \boxplus' by $\langle \Sigma_i \rangle'$, it is sufficient to prove that $E_n \cap (\langle \Sigma_1 \rangle' + \langle \Sigma_2 \rangle') \neq \emptyset$. Since all E_i as well as $[-1 - \lambda/\sqrt{n}, 1 + \lambda/\sqrt{n}]^n$ are Cartesian products, we may show this coordinatewise. For $C, D \subseteq \mathbb{R}$, let

$$C \boxplus'' D := (C + D) \cap [-1 - \lambda/\sqrt{n}, 1 + \lambda/\sqrt{n}],$$

and for any family Σ of subsets of \mathbb{R}^n , let $\pi_i(\Sigma)$ denote the family $\{\pi_i(A) : A \in \Sigma\}$ of subsets of \mathbb{R} , where $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ is the i th coordinate projection. We have to show the following:

$$1 + \lambda \in \langle \pi_n(\Sigma_1) \rangle'' + \langle \pi_n(\Sigma_2) \rangle'', \tag{10}$$

and for all $i \in [n-1]$,

$$[-1, 1] \cap \langle \pi_i(\Sigma) \rangle'' \neq \emptyset, \tag{11}$$

where the parenthesizations are with respect to \boxplus'' . First note the following.

CLAIM 1. *Any parenthesization with respect to \boxplus'' of one or more sets all equal to $[-1, 1]$ contains $[-1, 1]$.*

It follows from Claim 1 and induction that

CLAIM 2. *Any parenthesization with respect to \boxplus'' of two or more sets all but one equal to $[-1, 1]$, and the remaining set equal to $\{1 + \lambda\}$, contains $[\lambda, 1 + \lambda/\sqrt{n}]$.*

The interval $[\lambda, 1 + \lambda/\sqrt{n}]$ is non-empty since $\lambda \leq \sqrt{n}/(\sqrt{n} - 1)$ by hypothesis.

Since $\pi_n(\Sigma_1)$ consists of $[-1, 1]$'s (perhaps none) and a $\{1 + \lambda\}$, we obtain from Claim 2 that either $\langle \pi_n(\Sigma_1) \rangle'' \supseteq [\lambda, 1 + \lambda/\sqrt{n}]$ or $\langle \pi_n(\Sigma_1) \rangle'' = \{1 + \lambda\}$. Similarly, since $\pi_n(\Sigma_2)$ consists only of $[-1, 1]$'s, we have $\langle \pi_n(\Sigma_2) \rangle'' \supseteq [-1, 1]$ by Claim 1, and (10) follows.

Let $i \in [n - 1]$. Then $\pi_i(\Sigma)$ consists of a $\{1 + \lambda\}$, a $\{-1 - \lambda\}$, and $2n - 3$ $[-1, 1]$'s. The parenthesization $\langle \pi_i(\Sigma) \rangle''$ has a unique subexpression $\langle \pi_i(\Sigma_0) \rangle'' = \langle \pi_i(\Sigma_1) \rangle'' \boxplus'' \langle \pi_i(\Sigma_2) \rangle''$ such that $\{1 + \lambda\} \in \pi_i(\Sigma_1)$ and $\{-1 - \lambda\} \in \pi_i(\Sigma_2)$. By Claim 2, either $\langle \pi_i(\Sigma_1) \rangle'' \supseteq [\lambda, 1 + \lambda/\sqrt{n}]$ or $\langle \pi_i(\Sigma_1) \rangle'' = \{1 + \lambda\}$. Similarly, either $\langle \pi_i(\Sigma_2) \rangle'' \supseteq [-1 - \lambda/\sqrt{n}, -\lambda]$ or $\langle \pi_i(\Sigma_2) \rangle'' = \{-1 - \lambda\}$. This gives four cases, namely

- $\langle \pi_i(\Sigma_0) \rangle'' \supseteq [-1 + \lambda - \lambda/\sqrt{n}, 1 - \lambda + \lambda/\sqrt{n}]$,
- $\langle \pi_i(\Sigma_0) \rangle'' \supseteq [-1, -\lambda + \lambda/\sqrt{n}]$,
- $\langle \pi_i(\Sigma_0) \rangle'' \supseteq [\lambda - \lambda/\sqrt{n}, 1]$,
- $\langle \pi_i(\Sigma_0) \rangle'' \supseteq \{0\}$.

In all four cases it follows that $\langle \pi_i(\Sigma_0) \rangle'' \cap [-1, 1] \neq \emptyset$ (the first three cases because $\lambda \leq 1/(1 - \lambda/\sqrt{n})$), and we obtain (11) from the following claim, which is proved similarly to Claim 2.

CLAIM 3. *Any parenthesization with respect to \boxplus'' of two or more sets all but one equal to $[-1, 1]$, and the remaining set having non-empty intersection with $[-1, 1]$, has non-empty intersection with $[-1, 1]$.* \square

Acknowledgements

I would like to thank Nic van Rensburg for drawing my attention to the sub-differential calculus, as well as Frank Morgan and Mark Conger for their remarks and encouragement.

References

1. M. Alfaro, M. Conger, K. Hodges, A. Levy, R. Kochar, L. Kuklinski, Z. Mahmood, and K. von Haam, *The structure of singularities in Φ -minimizing networks in \mathbf{R}^2* , Pacific J. Math. **149** (1991), 201–210.
2. I. Anderson, *Combinatorics of Finite Sets*, Oxford University Press, Oxford, 1987.
3. A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. Ziegler, *Oriented Matroids*, 2nd ed., Encyclopedia of Mathematics and its Applications 46, Cambridge University Press, 1999.
4. H. Chen, C.-K. Cheng, A. B. Kahng, I. Măndoiu, Q. Wang, B. Yao, *The Y-architecture for on-chip interconnect: analysis and methodology*, Proc. Intern. Conf. CAD (ICCAD'03), ACM, 2003, pp. 13–19.
5. D. Cieslik, *Knotengrade kürzester Bäume in endlichdimensionalen Banachräumen*, Rostock Math. Kolloq. **39** (1990), 89–93.
6. D. Cieslik, *The vertex-degrees of Steiner minimal trees in Minkowski planes*, Topics in Combinatorics and Graph Theory (R. Bodendiek and R. Henn, eds.), Physica-Verlag, Heidelberg, 1990, pp. 201–206.
7. D. Cieslik, *Steiner minimal trees*, Nonconvex optimization and its applications, vol. 23, Kluwer, Dordrecht, 1998.

8. D. Cieslik, *Shortest connectivity*, Combinatorial Optimization, vol. 17, Springer, New York, 2005.
9. E. J. Cockayne, *On the Steiner problem*, Canad. Math. Bull. **10** (1967), 431–450.
10. M. Conger, Energy-Minimizing Networks in \mathbf{R}^n , Honours Thesis, Williams College, Williams-town MA, 1989.
11. A. de Bonis, G. O. H. Katona, K. J. Swanepoel, *Largest family without $A \cup B \subseteq C \cap D$* , J. Comb. Th., Ser. A **111** (2005), 331–336.
12. D.-Z. Du, B. Gao, R. L. Graham, Z.-C. Liu, and P.-J. Wan, *Minimum Steiner trees in normed planes*, Discrete Comput. Geom. **9** (1993), 351–370.
13. P. Erdős, *On a lemma of Littlewood and Offord*, Bull. Am. Math. Soc. **51** (1945), 898–902.
14. U. Fößmeier and M. Kaufmann, *Solving rectilinear Steiner tree problems exactly in theory and practice*, Proceedings of the fifth annual European symposium on algorithms (Gerhard Woeginger, ed.), Lecture Notes in Computer Science, vol. 1284, Springer-Verlag, 1997, pp. 171–185.
15. E. N. Gilbert and H. O. Pollak, *Steiner minimal trees*, SIAM J. Appl. Math. **16** (1968), 1–29.
16. H. Hadwiger, *Über Treffanzahlen bei translationsgleichen Eikörpern*, Arch. Math. **8** (1957), 212–213.
17. M. Hanan, *On Steiner's problem with rectilinear distance*, SIAM J. Appl. Math. **14** (1966), 255–265.
18. F. K. Hwang, D. S. Richards, and P. Winter, *The Steiner tree problem*, Ann. Discrete Math., vol. 53, North-Holland, Amsterdam, 1992.
19. A. O. Ivanov and A. A. Tuzhilin, *Minimal networks — the Steiner problem and its generalizations*, CRC Press, Boca Raton, FL, 1994.
20. B. Korte, H. J. Prömel, and A. Steger, *Steiner trees in VLSI-layout*, Paths, Flows, and VLSI-Layout (B. Korte et al., ed.), Algorithms and Combinatorics, vol. 9, Springer-Verlag, Berlin, 1990, pp. 185–214.
21. G. R. Lawlor and F. Morgan, *Paired calibrations applied to soap films, immiscible fluids, and surfaces and networks minimizing other norms*, Pacific J. Math. **166** (1994), 55–82.
22. D. T. Lee and C.-F. Shen, *The Steiner minimal tree problem in the λ -geometry plane*, Proceedings of the seventh international symposium on Algorithms and Computation (Tetsuo Asano et al., ed.), Lecture Notes in Computer Science, vol. 1178, Springer-Verlag, 1996, pp. 247–255.
23. D. T. Lee, C.-F. Shen, and C. L. Ding, *On Steiner tree problem with 45° routing*, 1995 International Conference on Circuits and Systems, IEEE, 1995, pp. 1680–1683.
24. Y. Y. Li, S. K. Cheung, K. S. Leung, and C. K. Wong, *Steiner tree constructions in λ_3 metric*, IEEE Trans. Circuits Systems II **45** (1998), 563–574.
25. G.-H. Lin and G. L. Xue, *The Steiner problem in λ_4 -geometry plane*, 1998 International Symposium on Algorithms and Computation, Lecture Notes in Computer Science, vol. 1533, Springer-Verlag, 1998, pp. 327–336.
26. G.-H. Lin, G. L. Xue, and D. Zhou, *Approximating hexagonal Steiner minimal trees by fast optimal layout of minimum spanning trees*, 1999 International Conference on Computer Design, IEEE, 1999, pp. 392–398.
27. F. Morgan, *The cone over the Clifford torus in \mathbf{R}^4 is Φ -minimizing*, Math. Ann. **289** (1991), 341–354.
28. F. Morgan, *Minimal surfaces, crystals, networks, and undergraduate research*, Math. Intelligencer **14** (1992), 37–44.
29. F. Morgan, *Riemannian Geometry, A Beginner's Guide*, 2nd ed., A. K. Peters, 1998.
30. G. Robins and J. S. Salowe, *Low-degree minimum spanning trees*, Discrete Comput. Geom. **14** (1995), 151–165.
31. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
32. O. Rodrigues, *Sur le nombre de manières déffectuer un produit de n facteurs*, J. Math. Pures Appl. (1) **3** (1838), 549.
33. E. Schröder, *Vier combinatorische Probleme*, Z. für Math. Phys. **15** (1870), 361–376.
34. A. Schürmann and K. J. Swanepoel, *Three-dimensional antipodal and norm-equilateral sets*, to appear in Pacific J. Math.
35. R. P. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge University Press, 1999.
36. K. J. Swanepoel, *Vertex degrees of Steiner Minimal Trees in ℓ_p^d and other smooth Minkowski spaces*, Discrete Comput. Geom. **21** (1999), 437–447.
37. K. J. Swanepoel, *The local Steiner problem in normed planes*, Networks **36** (2000), 104–113.

38. K. J. Swanepoel, *Quantitative illumination of convex bodies and vertex degrees of geometric Steiner minimal trees*, to appear in *Mathematika*.
39. A. C. Thompson, *Minkowski Geometry*, Encyclopedia of Mathematics and its Applications 63, Cambridge University Press, 1996.
40. C. Zălinescu, *Convex analysis in general vector spaces*, World Scientific, 2002.

DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF SOUTH AFRICA, PO BOX 392,
PRETORIA 0003, SOUTH AFRICA

E-mail address: swanekj@unisa.ac.za